University of Groningen

Phase transitions in soft-committee machines

Biehl, Michael; Ahr, Martin; Schlösser, Enno

Published in:
Computer Physics Communications

DOI:
[10.1016/S0010-4655\(06\)70021-2](https://doi.org/10.1016/S0010-4655(06)70021-2)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
1999

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):
Biehl, M., Ahr, M., & Schlösser, E. (1999). Phase transitions in soft-committee machines. *Computer Physics Communications*, 121-122, 614. [https://doi.org/10.1016/S0010-4655\(06\)70021-2](https://doi.org/10.1016/S0010-4655(06)70021-2)

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Abstract

Phase transitions in soft-committee machines

Michael Biehl¹, Martin Ahr, Enno Schlösser

Institut für Theoretische Physik, Universität Würzburg, Am Hubland, D-97074 Würzburg, Germany

Off-line learning in two-layered neural networks with continuous, sigmoidal activation functions in the hidden layer and a single linear output (soft-committee machines) is studied in a statistical physics framework. So far, such networks have only been investigated in situations where a sequence of independent training examples is provided (on-line learning).

Here we consider a student with K hidden units learning a rule which can be parametrized by a teacher of the same architecture. The training of the student is guided by the minimization of a quadratic error measure (training energy) with respect to a given set of examples.

The corresponding free energy, on average over random training examples, can be calculated analytically for arbitrary K in the high temperature limit. The system undergoes a phase transition from poor to good generalization when an increasing number of examples is presented. The transition is found to be second order for $K = 2$ and first order for all $K \geq 3$. Our results are supported by Monte Carlo simulations of the learning process. The behavior of large networks in the limit $K \rightarrow \infty$ is of particular interest. In this limit it is possible to extend the analysis to (more realistic) lower training temperatures. © 1999 Elsevier Science B.V. All rights reserved.

¹ E-mail: biehl@physik.uni-wuerzburg.de.